

# ajava

Asian Journal of Animal and Veterinary Advances



Academic  
Journals Inc.

[www.academicjournals.com](http://www.academicjournals.com)

## Prediction of Second Parity Milk Performance of Dairy Cows from First Parity Information Using Artificial Neural Network and Multiple Linear Regression Methods

<sup>1</sup>M.A. Edriss, <sup>1</sup>P. Hosseinnia, <sup>2</sup>M. Edrisi, <sup>1</sup>H.R. Rahmani and <sup>3</sup>M.A. Nilforooshan

<sup>1</sup>Department of Animal Science, Isfahan University of Technology, Isfahan, Iran

<sup>2</sup>Departments of IT and Biomedical Eng., University of Isfahan, Isfahan, Iran

<sup>3</sup>Department of Animal Breeding and Genetics,  
Swedish University of Agricultural Sciences, Uppsala, Sweden

---

**Abstract:** A mathematical model for prediction of second parity milk yield and fat percentage, with the use of first parity information seems to be helpful in order to predict the performance of prospective productive cows. As a tool for this prediction, back propagation neural network and multiple linear regression methods were compared based on their prediction differences with observed values. While, multiple linear regressions are based on linear relationships between variables, artificial neural network system also considers non-linear relationships between parameters. Data was collected from 4 medium sized dairy herds in Isfahan, Iran, which was divided into three parts in order to train, verify and test the artificial neural network system and estimation of regression coefficients, verify and test the multiple linear regression method. The results of the simulation showed that evaluations from both multiple linear regression and artificial neural network methods are good predictors for second parity production estimated from first parity information. However, artificial neural network predictions showed lower differences with the observed values and better quality parameters than multiple linear regression predictions, which made this assumption that artificial neural network system is more accurate in prediction.

**Key words:** Neural network, multiple linear regression, milk yield, fat percentage

---

### INTRODUCTION

Production traits in dairy cattle are under the influence of many genetic and environmental factors and the interactions between them both linearly and non-linearly. Dairy production traits including milk yield and milk fat percentage are of the most important economic traits in Iran's dairy industry. So, prediction of these traits is of importance to find prospective high yielding cows and improving the economic proficiency of dairy farms. Also, much of the selection of superior bulls is based on their ability to produce high yielding cows (Salehi *et al.*, 1998). Accuracy rate of finding high producing cows is important, because feeding, breeding, maintenance, veterinary and other costs can be saved for superiors and also by mis-culling cows of high genetic value, good sources of gene pool will be lost. In many countries, analysis of milk yield for 305 day lactation period is a foundation for dairy cattle genetic evaluations. So, implementing mathematical models for prediction of 305 day production in subsequent lactations from previous lactations or predicting total lactation yield from early records would be useful.

In comparison with regression methods or time-series analyses, Artificial Neural Network (ANN) represents a different new approach. The relationship between two or more independent

variables on a dependent variable can be obtained applying Multiple Linear Regression (MLR) method. Regressions show the extent and direction of associations between characters in the units of measurements. A MLR explains the linear cause-consequence relationships between some independent variables ( $x_1, x_2, \dots, x_n$ ) and a dependent variable ( $y$ ). Artificial Neural Network (ANN), like biological neural network, is made up from some sets of neurons. These neurons process the presented input and matching output to input in a supervised manner and extract non-linear relationships between those input and output. ANN consists of a set of neurons which are connected by weighted links that pass signals from one neuron to another. During training, the weights become adjusted to reduce error between actual and desired output. This error is minimized until it reaches to a certain objective value (Md Saad *et al.*, 2007).

ANN proposes an approach that is completely different from those offered by conventional methods. It solves particular problems through a learning system by typical inputs and specific desired outputs. (Grzesiak *et al.*, 2003). The usefulness of any mathematical model depends on how well it can mimic the biological process of milk production and adjusts for factors affecting it (Olori *et al.*, 1999). In ANN identifying patterns and relationships between the input and the corresponding output in a sample data set refers to the fact that optimal net performance depends on the recognition and extraction of non-linear relations through the training step which form the ANN structure (Lacroix *et al.*, 1995). Using this relation in the simulation stage, ANN can anticipate the output of the problem in a complex biological system from known input. In practice, ANNs have been primarily used in engineering, economics, or even in detection of heart abnormalities (Md Saad *et al.*, 2007). Recently, they also have been used in some areas of animal genetics and husbandry, such as detection of clinical disease (Yang *et al.*, 1999), estimating meat quality (Brethour, 1994), prediction of slaughter value of bulls (Adameczyk *et al.*, 2005), evaluation of physiological status of cows (Molenda *et al.*, 2001), detection of mastitis in dairy cattle (López-Benavide *et al.*, 2003), to predict swine daily gain in different ambient temperatures (Korthals *et al.*, 1994), prediction and classification of dairy cows based on milk yield in one period (Salehi *et al.*, 1998) and prediction of 305 day milk production from part lactation records (Lacroix *et al.*, 1995).

The aim of this research was comparing the predictive ability and the accuracy of ANN and MLR methods for predicting 305 day adjusted kilogram Milk Yield (MY) and Milk Fat Percentage (%MF) of the second lactation using information from first lactation as a tool for recognition of more producer cows of high genetic merit as the parents of the next generation.

## MATERIALS AND METHODS

The data was provided by the Animal Husbandry Division, Agricultural Organization of the Ministry of Agriculture in Isfahan, Iran, which was consisted of collected information from 32 Holstein dairy herds milked during 1995 to 2002. From the available herds, four medium sized herds were selected randomly for final investigations. Records were restricted to cows with completed second lactation. Followed by this restriction, a sample of 1880 cows with records was made available for further studies. The sample data was consisted of cows' registration number, purity (% Holstein blood which was  $65.5 \pm 19.43$  in the sample), first and second parities milk yield and fat percentage, corrected for 305 days in milk and some other information on the first parity of cows. Then, the data structure was rechecked and the data was introduced to MATLAB (2006) software for further processing.

Ten variables of the first parity (as inputs) plus two variables of 305 day MY and %MF from the second parity (as outputs) were assigned to each cow for both ANN and MLR (Table 1). In order to achieve a better learning for ANN, first lactation 305 day MY was classified into 9 production levels including milk production <2000 kg as the first and >9000 kg as the ninth level and the middle

Table 1: Variables used in the experimental data sets

Input data (1st parity)	Output data (2nd parity)
%Fat 305 day	Milk yield 305 day (kg)
Milk yield 305 day (kg)	%Fat 305 day
Cumulative fat yield (kg)	
Cumulative milk yield (kg)	
Milking times (2X, 3X)	
Average daily milk yield (kg)	
Purity*	
Calving season (3-mo)	
Number of records	
Age at calving	

\*: %Holstein blood

levels of 1000 kg difference. Salehi *et al.* (1998) concluded that data classification would lead to a better network learning. Normality distribution of each set was tested using Statistical Analysis Software (SAS, 1997). For ANN, the minimum and maximum values of each variable (Table 1) were mapped to the mean and standard deviation of 0 and 1, respectively. In order to construct the network, the neural network toolbox of MATLAB (2006) was used. The constructed network was a back propagation artificial neural network which had 3 layers of input, hidden and output with 10, 10 and 2 neurons in each, respectively. For input and hidden layers, tangent hyperbolic transfer function and for output layer, purline transfer function were used (MATLAB, 2006). The net learning function updated the weight and bias values relative to Levenberg-Marquardt optimization algorithm (Hagan and Menhaj, 1994). The net trained in 100,000 cycles of processing elements.

Cows were assigned to two groups:

**Group 1:** Included 1850 cows. Data of the second parity corresponding to the first period were used to design ANN and MLR for both MY and %MF. This part of data was divided into 925 training and 925 verification sets. The training set was used to obtain and modify the weights by ANN and to obtain the related regression coefficients by MLR. Verification set was used to control the size of network error during the training step and consequently to control the approximation ability of the network (Grzesiak *et al.*, 2006).

**Group 2:** Thirty cows were randomly selected from 1880 cows as a simulation set. The simulation set was used to test both ANN and MLR by predicting second parity MY and %MF from first parity information and then comparing the results of ANN anticipations and the results taken from MLR regression coefficients with the observed values.

MATLAB (2006) and SAS (1997) softwares were employed to run ANN and MLR analyses, respectively. The criterions used to compare the results of ANN and MLR anticipations with the actual observed data were: (1) adjusted coefficient of determination, (2) root of mean square error, (3)  $SD_{ratio}$ , (4) Pearson's coefficient of correlation between observed and predicted values, (5) relative mean error of prediction and (6) Theil's inequality coefficient.

$$R_A^2 = 1 - \frac{n-1}{n-k} \times (1 - R^2) \quad (1)$$

Where:

$R_A^2$  = Adjusted coefficient of determination

n = No. of records

k = No. of predictors or independent variables

$R^2$  = Coefficient of determination

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y})^2}{n}} \quad (2)$$

Where:

- RMSE = Root of mean square error
- n = No. of records
- $y_i$  = Observed value
- $\hat{y}$  = Estimated value by ANN or MLR

$$SD_{ratio} = \sqrt{\frac{\sum_{i=1}^n (e_i - \bar{e})^2}{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (3)$$

Where:

- $SD_{ratio}$  = Ratio of error standard deviation to the total standard deviation
- $e_i$  = Individual error
- $\bar{e}$  = Mean of error values
- $y_i$  = Observed value
- $\bar{y}$  = Mean of observed values

$$r_p = \delta_{ip} / \delta_i \delta_p \quad (4)$$

Where:

- $r_p$  = Pearson's correlation coefficient between observed and predicted values
- $\delta_{ip}$  = Covariance between observed and predicted values
- $\delta_i$  = Standard deviation of observed values
- $\delta_p$  = Standard deviation of predicted values

$$\Psi = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}}{y_i} \right| \times 100\% \quad (5)$$

Where,  $\Psi$  is the relative mean error of prediction and the other symbols are as the same as for the previous formulas.

$$I^2 = \frac{\sum_{i=1}^n (y_i - \hat{y})^2}{\sum_{i=1}^n y_i^2} \quad (6)$$

Where,  $I^2$  is Theil's inequality coefficient (Theil, 1979) and the other symbols are as the same as for the previous formulas.

The above coefficient is the sum of three other model's inequality coefficients.

$$I^2 = I_o^2 + I_B^2 + I_E^2 \quad (7)$$

The components of Eq. 7 are as follows:

$$I_o^2 = \frac{(\bar{y}_i - \hat{y}_m)^2}{(1/n) \sum_{i=1}^n y_i^2} \quad (8)$$

Where:

$I_o^2$  = Prediction bias

$\bar{y}_i$  = Mean of observed values

$\hat{y}_m$  = Mean of predicted values

$$I_B^2 = \frac{(\delta_i - \delta_p)^2}{(1/n) \sum_{i=1}^n y_i^2} \quad (9)$$

Where,  $I_B^2$  represents the error resulting from predictions' inadequate flexibility.

$$I_E^2 = \frac{2\delta_i \delta_p (1 - r_p)}{(1/n) \sum_{i=1}^n y_i^2} \quad (10)$$

Where,  $I_E^2$  represents the error resulting from insufficient convergency between direction of changes in the observed values and changes in the predicted values.

## RESULTS AND DISCUSSION

Regression coefficients estimated by MLR method for MY and %MF are shown in Table 2. These estimated regression coefficients obtained from the train set were used to make evaluations for the test set. ANN and MLR predictions were compared to the observed values by their mean differences to the mean of observed values (Table 3). Although MLR showed very low differences close to zero for train and verification data sets, finally, ANN predictions had lower differences with the observed data, may be due to the fact that MLR has no learning ability and it only finds linear relationships between data. The variability parameters (SD, CV) were also closer to those for observed data for ANN than MLR. Grzesiak *et al.* (2003) using test day records to estimate 305-d lactation yield, derived only 13.2 kg higher and 91.3 kg lower milk yield than the average of actual yield for ANN and MLR predictions, respectively.

Some quality parameters are shown for ANN and MLR for both MY and %MF by Table 4.  $SD_{ratio}$  was lower and  $R_A^2$  was higher for ANN relative to MLR. Also for ANN,  $SD_{ratio}$  decreased and  $R_A^2$  increased in the test step. These findings show the relative advantage of ANN to MLR. Better quality parameters for ANN relative to MLR have been also reported by Grzesiak *et al.* (2003, 2006). Regardless of the method of evaluation,  $SD_{ratio}$  and %RMSE were lower and  $R_A^2$  was higher for %MF relative to MY, which show that the input variables may better justified the changes in %MF than MY. Due to training and verification abilities of ANN, its quality of prediction drastically improved in the test step. However, it was not expected from MLR to show any considerable improvement in the test step relative to the previous steps.  $SD_{ratio}$  value less than 0.4 shows a good quality of the model, whereas values lower than 0.1 mean that the model would be close to ideal (Grzesiak *et al.*, 2006). In this study, low  $R_A^2$  values were derived for MLR, which showed that the chosen independent variables alone, could not explain well the changes in the dependent variable by MLR method.  $R_A^2 = 0.70$  implies a very good fitness for the model. While,  $R_A^2 < 0.40$  shows a

Table 2: Estimated regression coefficients using MLR method

	MY	%MF
Intercept	2244.04±1508.25	(1322.14±396.9)e-3
$\beta_1$	249.28±317.16	(454.72±83.46)e-3
$\beta_2$	0.15±0.20	(-0.07±0.05)e-3
$\beta_3$	0.44±3.65	(2.13±0.96)e-3
$\beta_4$	0.18±0.12	(-0.04±0.03)e-3
$\beta_5$	-3.26±2.98	(0.26±0.78)e-3
$\beta_6$	85.56±50.90	(13.74±13.40)e-3
$\beta_7$	70.83±27.98	(-12.19±7.36)e-3
$\beta_8$	-0.45±1.97	(2.45±0.52)e-3
$\beta_9$	-207.45±114.39	(38.47±30.10)e-3
$\beta_{10}$	154.83±30.13	(-23.09±7.93)e-3

MY: Milk Yield; %MF: Milk Fat Percentage, Correlation coefficients ( $\beta$ ) are in the same sequence as the input variables represented in Table 1

Table 3: Descriptive parameters for the observed and predicted (by ANN and MLR) data

Parameters	OBS			ANN			MLR			
	Train	Verify	Test	Train	Verify	Test	Train	Verify	Test	
MY	Mean	7471.37	7522.21	7342.98	7485.37	7482.72	7400.70	7471.38	7522.23	7473.44
	DIFF	-	-	-	14.00	-39.50	57.72	0.01	0.02	130.46
	t	-	-	-	-0.26	0.71	-0.20	0.00	0.00	-0.46
	SD	1400.69	1434.25	1348.01	907.61	899.78	784.35	846.52	872.83	761.95
	CV	18.74	19.06	18.35	12.12	12.02	10.59	11.33	11.60	10.19
%MF	Mean	2.73	2.70	2.67	2.73	2.72	2.74	2.73	2.70	2.72
	DIFF	-	-	-	0.00	0.02	0.07	0.00	0.00	0.05
	t	-	-	-	0.00	-1.65	-0.59	0.00	0.00	-0.38
	SD	0.39	0.37	0.56	0.27	0.26	0.27	0.25	0.24	0.26
	CV	14.27	13.78	20.84	9.96	9.66	9.88	9.39	8.84	9.59

MY: Milk yield; %MF: Milk fat percentage; OBS: Observed value; ANN: Artificial neural network prediction; MLR: Multiple linear regression prediction; DIFF: Difference from the mean of observed values; t: All DIFF's have no significant difference from zero ( $p > 0.05$ )

Table 4: Quality parameters for ANN and MLR methods

Parameters	MY				%MF				
	$SD_{ratio}$	$R^2_A$	RMSE	%RMSE	$SD_{ratio}$	$R^2_A$	RMSE	%RMSE	
ANN	Train	0.754	0.425	1015.16	14.1	0.710	0.489	0.276	10.1
	Verify	0.820	0.324	1172.74	15.6	0.796	0.367	0.297	10.9
	Test	0.576	0.637	817.48	10.9	0.106	0.729	0.336	12.6
MLR	Train	0.797	0.358	1115.37	14.9	0.753	0.425	0.293	10.7
	Verify	0.998	0.366	1137.47	15.1	0.770	0.403	0.280	10.3
	Test	0.697	0.327	933.93	12.5	0.544	0.403	0.350	12.8

MY: Milk yield; %MF: Milk fat percentage;  $SD_{ratio}$ : The ratio of error standard deviation to the total standard deviation;  $R^2_A$ : Adjusted coefficient of determination; RMSE: Root of mean square error; %RMSE: RMSE divided by the mean of performance

non-appropriate model (Olori *et al.*, 1999). Although,  $R^2_A$  was low in the verification set, the final predictions by ANN for the test data set had a high  $R^2_A$ . In other studies which have used partial records to predict full lactation records, higher  $R^2_A$  values were estimated, including  $R^2_A = 0.79$  by Wood (1967) and  $R^2_A = 0.94$  by Olori *et al.* (1999). Grzesiak *et al.* (2003) reported RMSEs equal to 501.7 and 544.76 kg milk yield in the test step for ANN and MLR, respectively. Also, Salehi *et al.* (1998) estimated RMSE values ranging from 445 to 554 kg depending on the network system and the average of herd milk production. The reason for the differences between the results of these studies and the results of the current study refers to different data structures used to train ANN. For example, they have used test day records to calculate 305 day milk production with more or less extended data from other regions with more input variables.

Table 5: Predictive measures for ANN and MLR

Parameters		$r_p$	$\Psi$	$I^2$	$I^2_O$	$I^2_B$	$I^2_E$
MY	ANN	Train	0.657	11.74	0.019	3.4e-6	4.2e-3
		Verify	0.577	12.57	0.023	2.0e-11	4.8e-3
		Test	0.825	10.13	0.012	5.9e-5	5.7e-3
	MLR	Train	0.604	12.38	0.021	2.0e-11	5.3e-3
		Verify	0.608	12.36	0.022	2.0e-11	5.4e-3
		Test	0.737	11.90	0.016	3.1e-4	6.2e-3
%MF	ANN	Train	0.703	8.03	0.010	6.0e-11	18.0e-5
		Verify	0.612	8.72	0.011	8.2e-5	16.0e-4
		Test	0.901	9.80	0.015	6.0e-4	11.0e-3
	MLR	Train	0.657	8.50	0.011	7.5e-4	23.4e-4
		Verify	0.410	8.28	0.011	1.0e-14	23.9e-4
		Test	0.870	9.97	0.016	2.5e-4	117.7e-4

MY: Milk yield; %MF: Milk fat percentage;  $r_p$ : Correlation coefficient with the observed data ( $p < 0.001$ );  $\Psi$ : Relative mean error of prediction;  $I^2$ : Theil's inequality coefficient;  $I^2_O$ : Prediction bias;  $I^2_B$ : Prediction inflexibility;  $I^2_E$ : Insufficient convergency between direction of changes in the observed and predicted values

Table 5 shows some parameters related to predictive ability of ANN and MLR for both MY and %MF. As shown in this table, on average,  $r_p$  values were higher and  $\Psi$  and  $I^2$  values were lower for ANN relative to MLR, which were in favor of ANN. These values showed a same situation for %MF, which support the results of Table 4 regarding better fitness of both ANN and MLR for %MF relative to MY. For both ANN and MLR,  $r_p$  increased for the test data set. Both  $\Psi$  and  $I^2$  decreased in the test step for MY, while they increased for %MF.

Considering final predictions obtained in the test step (Table 5), except  $I^2_O$  for %MF, all of the three criterions of  $I^2_O$ ,  $I^2_B$  and  $I^2_E$  were lower in both MY and %MF for ANN relative to MLR, which shows that ANN predictions are less bias and more flexible and the direction of changes in the observed and predicted data are more in convergence for ANN than MLR.

The most important part of Theil's coefficient ( $I^2$ ) was related to  $I^2_E$ , which represents an error resulting from a lack of full convergency in the direction of changes between the observed and predicted values, particularly for the neural network. This result was in agreement with the results obtained by Grzesiak *et al.* (2006).

The major use of any predictive process is to support accurate decisions which are dependent on a prior knowledge to make possible outcome(s). The results of this study showed that both MLR and ANN can be used to predict second parity production from first parity information. MLR models are simple to design and define parameters. However, the results showed that ANN systems have the ability to predict second parity 305 day milk yield and fat percentage with a higher accuracy. Correlations between the observed values and predictions, together with the other quality parameters and predictive measures had better situations for ANN relative to MLR. Also, ANN predictions showed lower deviations from the observed data, but this difference between ANN and MLR was slight and can be negligible. Adding new data requires a new statistical model, whereas a neural network system can update itself with new data. Finally, ANN can be improved with more additional input variables and training with more actual data to get more accurate predictions.

#### ACKNOWLEDGMENT

The authors would like to acknowledge Isfahan University of Technology for supporting financially P. Hosseinia's M.Sc. Thesis.

#### REFERENCES

Adamczyk, K., K. Molenda, J. Szarek and G. Skrzyński, 2005. Prediction of bulls slaughter from growth data using artificial neural network. *J. Cent. Eur. Agric.*, 6 (2): 133-142.



- Brethour, J.R., 1994. Estimating marbling score in live cattle from ultrasound images using pattern recognition and neural network procedures. *J. Anim. Sci.*, 72: 1425-1432.
- Grzesiak, W., R. Lacroix, J. Wójcik and P. Blaszczyk, 2003. A comparison of neural network and multiple regression predictions for 305 day lactation yield using partial lactation records. *Can. J. Anim. Sci.*, 83 (2): 307-310.
- Grzesiak, W., P. Blaszczyk and R. Lacroix, 2006. Methods of predicting milk yield in dairy cows- Predictive capabilities of Wood's lactation curve and artificial neural networks (ANNs). *Comput. Electron. Agric.*, 54: 69-83.
- Hagan, M.T. and M. Menhaj, 1994. Training feed forward networks with the Marquardt algorithm. *IEEE Trans. Neural Networks*, 5: 989-993.
- Korthals, R.L., G.L. Hahn and J.A. Nienaber, 1994. Evaluation of neural networks as a tool for management of swine environments. *Trans. ASAE*, 37: 1295-1299.
- Lacroix, R., K.M. Wade, R. Kok and J.F. Hayes, 1995. Prediction of cow performance with a connectionist model. *Trans. ASAE*, 38: 1573-1579.
- López-Benavide, M.G., S. Samarasinghe and J.G.H. Hickford, 2003. The use of artificial neural networks to diagnose mastitis in dairy cattle. *The International Joint Conference on Neural Networks (IJCNN)*: (CD-ROM).
- MATLAB, 2006. *MATLAB Guide for Personal Computers. Version 7.3 Edn.* MATLAB Institute, SA.
- Md Saad, M.H., M.J. Mohd Nor, F.R.A. Bustami and R. Ngadiran, 2007. Classification of heart abnormalities using artificial neural network. *J. Applied Sci.*, 7 (6): 820-825.
- Molenda, K., Z. Gil and J. Ychlińska, 2001. Conception of computer system support in detection of pregnancy, oestrus and mastitis based on milk temperature. *52nd Annual Meeting of EAAP, Budapest, Hungary.*
- Olori, V.E., S. Brotherstone, W.G. Hill and B.J. McGuirk, 1999. Fit of standard models of the lactation curve to weekly records of milk production of cows in a single herd. *Livest. Prod. Sci.*, 58: 55-63.
- Salehi, F., R. Lacroix and K.M. Wade, 1998. Effects of learning parameters and data presentation on the performance of back-propagation networks for milk yield prediction. *Trans. ASAE*, 41: 253-259.
- SAS, 1997. *User's Guide: Statistics, Version 6.12 Edn.* SAS Inst., Inc., Cary, NC.
- Theil, H., 1979. World income inequality. *Econ. Lett.*, 2: 99-102.
- Wood, P.D.P., 1967. Algebraic model of the lactation curve in cattle. *Nature*, 216: 164-165.
- Yang, X.Z., R. Lacroix and K.M. Wade, 1999. Neural detection of mastitis from dairy herd improvement records. *Trans. ASAE*, 42: 1063-1072.